



## Retina-inspired Filter

Effrosyni Doutsis, Lionel Fillatre, Marc Antonini, Julien Gaulmin

### ► To cite this version:

Effrosyni Doutsis, Lionel Fillatre, Marc Antonini, Julien Gaulmin. Retina-inspired Filter. IEEE Transactions on Image Processing, 2018, 27 (7), pp.3484 - 3499. 10.1109/TIP.2018.2812079 . hal-01350686

**HAL Id: hal-01350686**

**<https://hal.science/hal-01350686>**

Submitted on 1 Aug 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Retina-inspired Filter

Effrosyni Doutsis, *Member, IEEE*, Lionel Fillatre, Marc Antonini, *Member, IEEE*, and Julien Gaulmin

**Abstract**—This paper introduces a novel filter which is inspired by the human retina. The human retina consists of three different layers: the Outer Plexiform Layer (OPL), the inner plexiform layer and the ganglionic layer. Our inspiration is the linear transform which takes place in the OPL and has been mathematically described by the neuroscientific model “virtual retina”. This model is the cornerstone to derive the non-separable spatiotemporal OPL retina-inspired filter, briefly renamed retina-inspired filter, studied in this paper. This filter is connected to the dynamic behavior of the retina, which enables the retina to increase the sharpness of the visual stimulus during filtering before its transmission to the brain. We establish that this retina-inspired transform forms a group of spatiotemporal Weighted Difference of Gaussian (WDoG) filters when it is applied to a still image visible for a given time. We analyze the spatial frequency bandwidth of the retina-inspired filter with respect to time. It is shown that the WDoG spectrum varies from a lowpass filter to a bandpass filter. Therefore, while time increases, the retina-inspired filter enables to extract different kind of information from the input image. Finally, we discuss the benefits of using the retina-inspired filter in image processing.

**Index Terms**—Human retina, OPL layer, Spatiotemporal filtering, Weighted difference of gaussian.

## I. INTRODUCTION

THE decryption of the Human Visual System (HVS) is of a great interest. Studying its physiological and functional properties is helpful to many scientific fields, including signal and image processing. There are several bio-inspired filters for image analysis which have been built in order to mimic the HVS and the visual perception. Many filters are used as an approximation of the early vision system like static Difference of Gaussian (DoG), Gabor or foveating filters. This paper is focused on DoG filters which have been extensively studied in the past [1]–[4]. Most of the DoGs are static spatial filters which have been proposed for instance for perceptual visual quality metrics like the image quality assessment [5], for edge detection [6]–[8], and for object recognition [9]. Some works have studied multiscale DoG filters, introducing time as the scale level, in order to approximate the spatiotemporal behavior of the visual system. Laplacian pyramid [10], [11], visual saliency [12], [13], speed and direction estimation [14],

image stylization [4] and bio-inspired codecs [3], [15] are some of these attempts. However, the spatiotemporal receptive fields of the early HVS are more complex and above all dynamic. Wohrer *et al* [16] have proposed a more realistic model of the retina based on neuroscientific experiments. This paper proposes and studies a dynamic spatiotemporal DoG filter which is derived from [16] and adapted to the image processing field.

Our filter, which is inspired from a realistic model of the Outer Plexiform Layer (OPL), is named the non-separable spatiotemporal OPL retina-inspired filter but it is briefly called the retina-inspired filter. We have already briefly discussed about it in our previous works [17], [18]. The contributions of this paper are the followings. We first review the main neuroscientific models which have been proposed in order to mimic the linear transform which takes place in the OPL layer and the first group of cells of the retinal tissue [2], [16], [19]–[21]. This review underlines the importance of time for describing the way the retina cells are structured and work. Then, we derive from these previous works a realistic but simple retina-inspired filter. We show that this can be written as a Weighted DoG (WDoG) filter whose weights evolve in time when it is applied to a still image. Finally, we propose a full analysis of the WDoG filter in space and in frequency domains. We prove that its spatial frequency bandwidth depends on time, which means that while time increases the retina-inspired filter extracts different kind of information. More precisely, in the beginning, our filter is lowpass and it turns into bandpass while time increases. We also show that the filter stops to evolve asymptotically. As a result, there exists a certain time when the filter has already extracted all the necessary data.

In section II, we introduce the structure of the retina and we focus on the OPL on which the retina-inspired transform is based. In section III, we build the retina-inspired filtering model and study how it is applied to a still image. In section IV, we study the spatial and frequential behaviors of the retina-inspired filter considered as a group of WDoG filters. In section V, we illustrate the behavior of the filter with some numerical results. Last but not least, in section VI, we discuss the applications of this paper and some possible extensions of this work.

## II. NEUROSCIENTIFIC BACKGROUND: OPL MODEL

The visual system [22] is part of the Central Nervous System (CNS) and consists of many different areas which participate to the coding of a visual stimulus. The most important and better studied areas are the retina, the optic nerve, the Lateral Geniculate Nucleus (LGN) and the visual cortex as depicted in Fig. 1. The retina is a layer of tissue, lining the

E. Doutsis is with Université Côte d’Azur, CNRS, I3S, MediaCoding Team, UMR 7271, 06900 Sophia Antipolis, France and 4G-TECHNOLOGY, 460 avenue de la Quiera 06370 Mouans Sartoux - France.  
E-mail: see <http://doutsiefrosini.wix.com/effrosynidoutsis>

L. Fillatre is with Université Côte d’Azur, CNRS, I3S, MediaCoding Team, UMR 7271, 06900 Sophia Antipolis, France  
E-mail: [lionel.fillatre@i3s.unice.fr](mailto:lionel.fillatre@i3s.unice.fr)

M. Antonini is with Université Côte d’Azur, CNRS, I3S, MediaCoding Team, UMR 7271, 06900 Sophia Antipolis, France  
E-mail: [am@i3s.unice.fr](mailto:am@i3s.unice.fr)

J. Gaulmin is with 4G-TECHNOLOGY, 460 avenue de la Quiera 06370 Mouans Sartoux - France  
Site: <http://www.4g-technology.fr>

Thanks to 4G-TECHNOLOGY and ANRT.

inner surface of the eye, which is responsible to capture the visual stimuli and transform it into a sequence of electrical impulses (spike trains). This code of spikes is transmitted through the optic nerve to the LGN cells and finally to the primary visual cortex. The cells in visual cortex are more complex [23], [24] and their functions are related to edge and orientation detection, motion estimation, discrimination of the shape or the color, etc.

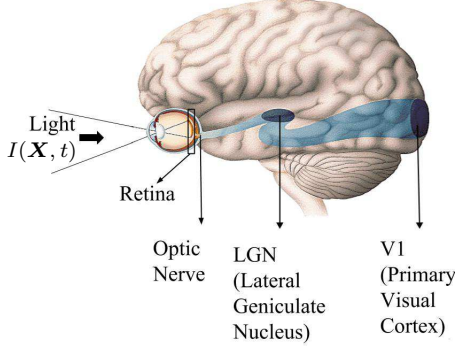


Fig. 1: The visual system pathway: the retina captures the visual stimuli  $I(\mathbf{X}, t)$  and transmits it to the visual cortex via the optic nerve and the LGN.

The input of the HVS is the luminance of light which is modeled as a 3D spatiotemporally varying visual stimulus  $I(\mathbf{X}, t)$  with  $\mathbf{X} \in \mathbb{R}^3$  and  $t \in \mathbb{R}$ . The luminance is transformed inside the *retina* [25]–[28] to fit the human brain. This transformation is serial due to the retina's structure. The retinal neurons can be separated into three groups: the Outer Plexiform Layer (OPL), the Inner Plexiform Layer (IPL) and the Ganglionic Layer (GL) as depicted in Fig. 2. The OPL layer is structured by photo-receptors, horizontal and bipolar cells. Bipolar and amacrine cells build the IPL layer. The GL layer is constructed by the amacrine and the ganglion cells [29], [30]. The overlapping of the layers is due to the feedforward and feedback network of the retinal cells. The OPL layer captures the incoming light (photons) and linearly transforms it into an electrical current. The IPL layer is responsible for a nonlinear rectification of the electrical signal. The GL produces the spike trains which is a kind of code based on electrical impulses [16].

This paper is dedicated to the study and analysis of the OPL layer in order to build a novel retina-inspired filter. According to the law of optics, the 3D visual stimuli  $I(\mathbf{X}, t)$  is projected onto the retina via the lens (the optics of the eye is detailed in [31]). Hence, the 3D luminance  $I(\mathbf{X}, t)$  is simplified into a 2D luminance  $f(\mathbf{x}, t)$  where  $\mathbf{x} \in \mathbb{R}^2$ . We assume that the visual stimulus is flashed for a given time  $T > 0$  and it is constant during this time interval. This involves that the 2D visual stimulus is written as  $f(\mathbf{x}, t) = f(\mathbf{x})\mathbb{1}_{[0, T]}(t)$  where  $f(\mathbf{x}) \in L^1(\mathbb{R}^2)$  is a still image and  $\mathbb{1}$  is the indicator function such that  $\mathbb{1}_{[0, T]}(t) = 1$  if  $0 \leq t \leq T$ , otherwise 0. The space  $L^1(\mathbb{R}^2)$  is the set of the Lebesgue integrable functions from  $\mathbb{R}^2$  to  $\mathbb{R}$ .

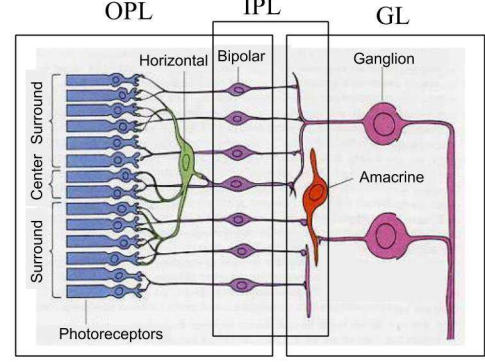


Fig. 2: The retinal layers according to [16]. This figure shows the connectivity and hierarchical structure of the retinal cells. It is important to note that cells of the same kind lie within a common layer.

#### A. OPL Layer

The OPL cells receive as an input the visual stimulus  $f(\mathbf{x}, t)$  which is spatiotemporally transformed into an electrical signal. This transformation takes place inside the *receptive field* (RF) of each cell. Kuffler [19] proposed to shape the RF by two concentric nested circles or ellipses, which are termed center-surround. The smaller circle (or ellipse) corresponds to the center and the larger one is the surround. Let  $\Omega_i \subseteq \mathbb{R}^2$  be the RF of a bipolar cell centered in  $\mathbf{x}_i \in \Omega_i$ . Let  $A(\mathbf{x}_i, t)$  be the electrical signal produced by the RF  $\Omega_i$  when the input signal is  $f(\mathbf{x}, t)$ . Considering the bipolar cell is a time invariant linear system, the following linear approximation of the OPL retinal-transform has been introduced in [29]:

$$A(\mathbf{x}_i, t) = \int_{t'=0}^{+\infty} \int_{\mathbf{x}' \in \Omega_i} K(\mathbf{x}_i - \mathbf{x}', t - t') f(\mathbf{x}', t') d\mathbf{x}' dt' \quad (1)$$

where  $K(\mathbf{x}, t)$  is the spatiotemporal transform of a single bipolar cell, also known as the Point Spread Function (PSF) at time  $t$ . The above equation indicates that the electrical signal  $A(\mathbf{x}_i, t)$  depends linearly on the spatial neighborhood and the past values of the input stimuli located in the RF  $\Omega_i$  of the single bipolar cell centered in  $\mathbf{x}_i$ . Assuming that i) the number of cells is very large, ii) all the cells obey to the same spatiotemporal model (spatial invariance) and iii) the temporal point spread function  $K(\mathbf{x}, t)$  is not restricted to the domain  $\Omega_i$ , the spatiotemporal transform (1) is approximated by the following spatiotemporal convolution denoted  $\overset{x, t}{*}$ :

$$\begin{aligned} A(\mathbf{x}, t) &= \int_{t' \in \mathbb{R}} \int_{\mathbf{x}' \in \mathbb{R}^2} K(\mathbf{x} - \mathbf{x}', t - t') f(\mathbf{x}', t') d\mathbf{x}' dt' \\ &= (K \overset{x, t}{*} f)(\mathbf{x}, t), \end{aligned} \quad (2)$$

where  $\mathbf{x} \in \mathbb{R}^2$  is located on the retina. Many models of the temporal point spread function  $K(\mathbf{x}, t)$  have been proposed.

#### B. Approximations of $K(\mathbf{x}, t)$

1) *Spatial Receptive Field*: The first mathematical approximation of the OPL was proposed by Kuffler [19]. A bipolar cell receives its input signal directly from a group of photoreceptors and/or a group of horizontal cells (Fig. 3). On the one hand, the output of two or more photoreceptors is averaged and transmitted to the center of the RF of the bipolar cell in order to excite it. This is approximated by a Gaussian filter  $G_{\sigma_c}(\mathbf{x})$  given by:

$$G_{\sigma_c}(\mathbf{x}) = \frac{1}{2\pi\sigma_c^2} \exp\left(-\frac{\|\mathbf{x}\|^2}{2\sigma_c^2}\right) = G_{\sigma_c}(\|\mathbf{x}\|), \quad (3)$$

where  $\mathbf{x} \in \mathbb{R}^2$ ,  $\|\mathbf{x}\|$  is the Euclidean norm of  $\mathbf{x}$  and  $\sigma_c$  is the standard deviation which tunes the spread of the Gaussian filter. On the other hand, the same or higher number of photoreceptors is linked to horizontal cells. An horizontal cell is strongly connected to neighbor horizontal cells averaging twice the initial input stimulus. The output of one or more horizontal cells is then propagated to the surround of the RF of the bipolar cell in order to inhibit it [32]. This signal is modeled by the Gaussian filter  $G_{\sigma_s}(\mathbf{x})$  with  $\sigma_c < \sigma_s$ . As a result, the bipolar cell receives two signals of opposite signs. Finally, the center-surround activity  $K(\mathbf{x}, t)$  of the RF of a bipolar cell is modeled as a DoG filter:

$$DoG(\mathbf{x}) = G_{\sigma_c}(\mathbf{x}) - G_{\sigma_s}(\mathbf{x}).$$

Kuffler assumed that all these processes happen instantaneously. Hence, the PSF  $K(\mathbf{x}, t)$  is constant for all time  $t$ .

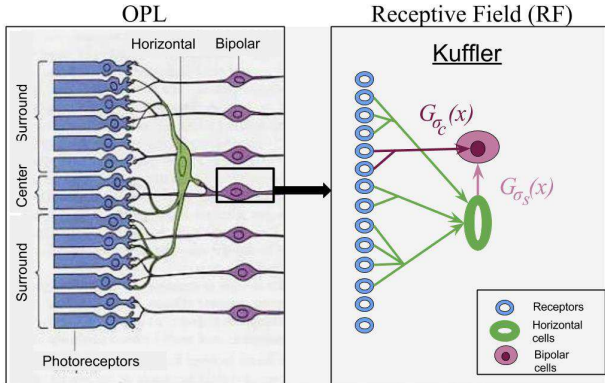


Fig. 3: Propagation of the electrical signal to the receptive field of a bipolar cell according to [19]. The left figure is part of the retinal structure which corresponds to the OPL layer. The right figure is focused on a single bipolar cell and how the RF of this cell works.

2) *Separable Spatiotemporal Receptive Field*: The first attempt to improve the static DoG filter was done by Marr [20]. Marr's theory contains an assumption of spatiotemporal separability adapting a temporal impulse response for each one of the decomposition layers which are formed by a static DoG:

$$K(\mathbf{x}, t) = H(t)DoG(\mathbf{x}), \quad (4)$$

where  $H(t)$  is usually low-pass for "sustained" units, and band-pass for "transient" units.

3) *Non-Separable Spatiotemporal Receptive Field*: The models above were not accurate with time-varying stimuli  $f(\mathbf{x}, t)$ . Hence, Fleet [21] proposed a non-separable spatiotemporal center-surround model as an extension of the DoG. Electrophysiological studies have shown that the center and the surround have different time courses of response. In addition, the temporal delay between the response of the center and surround areas of the cells receptive field should also be considered. This model was interpreted by Fleet as a precursor to the extraction of velocity specific information [21].

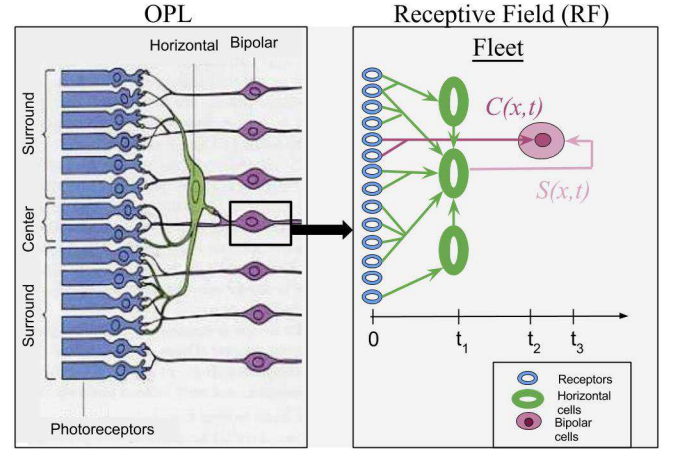


Fig. 4: Propagation of the electrical signal to the receptive field of a bipolar cell according to Fleet [21].

The inseparability of space and time was highlighted and confirmed while studying the spectrum of the center-surround model. The dynamics of the model are due to the sensitivity changes of the response of the cells and the different phases of the center-surround areas [21]. Similar models have been proposed in order to describe how does the receptive field of neurons work in areas which come after the retina like the Lateral Geniculate Nucleus (LGN) or cortical neurons [2], [29], [33]. The common point of all these models is the spatiotemporal inseparability, which confirms its importance.

A non-separable spatiotemporal retinal filtering as part of the visual system coding process has been mathematically introduced in *virtual retina* [29]. This filter  $K(\mathbf{x}, t)$ , given in (5), is built by the combination of Gaussian spatial filters and *Difference of Exponential (DoE)* temporal filters, where the center  $C(\mathbf{x}, t)$  and surround  $S(\mathbf{x}, t)$  areas are structured according to (6) and (7) respectively:

$$K(\mathbf{x}, t) = C(\mathbf{x}, t) - S(\mathbf{x}, t), \quad (5)$$

$$C(\mathbf{x}, t) = w_c G_{\sigma_c}(\mathbf{x}) V(t), \quad (6)$$

$$S(\mathbf{x}, t) = w_s G_{\sigma_s}(\mathbf{x}) \left( V^t * E_{\tau_s} \right)(t), \quad (7)$$

where  $w_c$  and  $w_s$  are constant parameters,  $G_{\sigma_c}(\mathbf{x})$  and  $G_{\sigma_s}(\mathbf{x})$  are spatial Gaussian filters standing for the center and surround areas respectively,  $V(t)$  is a low-pass filter and  $E_{\tau_s}(t)$  is an exponential temporal filter. The temporal filter  $V(t)$  is given by (8) and describes the *DoE* which stands for the spatial

variation with respect to time. It is modeled with temporal low-pass filters [16]:

$$V(t) = \begin{cases} E_{\tau_G, n} \ast (\delta_0 - w_c E_{\tau_C})(t) & \text{if } t \geq 0, \\ 0 & \text{otherwise,} \end{cases} \quad (8)$$

where the gamma temporal filter  $E_{\tau_G, n}(t)$  is defined by:

$$E_{\tau, n}(t) = \begin{cases} \frac{t^n \exp(-t/\tau)}{\tau^{n+1}} & \text{if } t \geq 0, \\ 0 & \text{otherwise,} \end{cases} \quad (9)$$

with  $n \in \mathbb{N}$ ,  $\tau > 0$ ,  $\delta_0(t)$  is the dirac function,  $E_{\tau_C}(t)$  is an exponential temporal filter, and  $\ast$  stands for the temporal convolution. The exponential temporal filter is given by (25) for  $n = 0$ .

The mathematical model of  $C(\mathbf{x}, t)$  and  $S(\mathbf{x}, t)$  is totally linked to the way the retinal cells are connected and propagate the stimuli. To explain the non-separable spatiotemporal retina-inspired filter, we need to focus again on bipolar cells. We have already mentioned how they receive two opposite signals in their RF. In this model the key is the hierarchy of the retinal cells and their connectivity. The temporal behavior of the retina cells is described in Fig. 4. The horizontal cells are the first ones which receive an average signal by a group of photoreceptors at time  $t_1$ . A smaller group of photoreceptors propagates an excitatory signal to the center of bipolar cells  $C(\mathbf{x}, t)$  at time  $t_2$  while horizontal cells receive this signal and they communicate and exchange information with adjacent horizontal cells. This causes a small delay  $E_{\tau_s}(t)$  until time  $t_3$  in the propagation of the inhibitory signal  $S(\mathbf{x}, t)$  coming from horizontal cells to bipolar cells.

### III. RETINA-INSPIRED FILTERING OF A STILL IMAGE

The goal of this section is to simplify the retinal spatiotemporal filtering (2), based on the model (5), when it is applied to a still image. We show that the retina-inspired filter could be interpreted as a group of time varying WDoG filters. That means that, at each time, a new and different spatial WDoG filter arises. Consequently, while time increases, the retina-inspired filter is able to extract different kinds of information from its input signal.

#### A. Equivalence with a Group of Time-varying WDoG Filters

A still image does not vary in time when it is flashed. This time invariance enables us to simplify the spatiotemporal convolution (2) as established in Proposition 1.

**Proposition 1.** Assume  $f(\mathbf{x}, t) = f(\mathbf{x}) \mathbb{1}_{[0, T]}(t)$  for all  $\mathbf{x} \in \mathbb{R}^2$  and all  $t \in \mathbb{R}$ . Then, the spatiotemporal convolution (2) turns into the spatial convolution:

$$A(\mathbf{x}, t) = \phi(\mathbf{x}, t) \ast f(\mathbf{x}), \quad (10)$$

where  $\phi(\mathbf{x}, t)$  is a spatiotemporal WDoG filter weighted by two temporal filters  $R_c(t)$  and  $R_s(t)$ :

$$\phi(\mathbf{x}, t) = \begin{cases} w_c R_c(t) G_{\sigma_c}(\mathbf{x}) - w_s R_s(t) G_{\sigma_s}(\mathbf{x}) & \text{if } t \geq 0, \\ 0 & \text{otherwise,} \end{cases} \quad (11)$$

$$R_c(t) = \mathbb{1}_{[0, +\infty)}(t) \int_{\max\{0, t-T\}}^t V(u) du, \quad (12)$$

$$R_s(t) = \mathbb{1}_{[0, +\infty)}(t) \int_{\max\{0, t-T\}}^t (V \ast E_{\tau_s})(u) du, \quad (13)$$

for all  $\mathbf{x} \in \mathbb{R}^2$  and for all  $t \in \mathbb{R}$ .

*Proof.* The proof consists in calculating the convolution:

$$\begin{aligned} A(\mathbf{x}, t) &= K(\mathbf{x}, t) \ast^{x, t} f(\mathbf{x}, t) \\ &= (C(\mathbf{x}, t) - S(\mathbf{x}, t)) \ast^{x, t} f(\mathbf{x}) \mathbb{1}_{[0, T]}(t) \\ &= w_c G_{\sigma_c}(\mathbf{x}) \ast^{x, t} f(\mathbf{x}) \left( V \ast \mathbb{1}_{[0, T]}(t) \right) \\ &\quad - w_s G_{\sigma_s}(\mathbf{x}) \ast^{x, t} f(\mathbf{x}) \left( \left( V \ast E_{\tau_s} \right) \ast \mathbb{1}_{[0, T]}(t) \right). \end{aligned}$$

For an integrable function  $U(t)$ , a short calculation shows that

$$U \ast \mathbb{1}_{[0, T]}(t) = \begin{cases} 0 & \text{if } t < 0, \\ \int_{\max\{0, t-T\}}^t U(u) du & \text{otherwise.} \end{cases} \quad (14)$$

Then, it is straightforward to derive (10) with  $\phi(\mathbf{x}, t)$  defined in (11).  $\square$

This proposition is fundamental because it turns spatio-temporal filtering of the still image  $f(\mathbf{x})$  with the retinal filter  $K(\mathbf{x}, t)$  into a simpler spatial convolution with the retina-inspired filter  $\phi(\mathbf{x}, t)$  which is a group of time-varying WDoGs. The temporal functions  $R_c(t)$  and  $R_s(t)$  act like weights and modify the WDoG spatial spectrum with respect to time (see Section IV). Since a DoG filter is a 2D spatially symmetric real function, the filter  $\phi(\mathbf{x}, t)$  can be simplified into  $\phi(r, t)$  where the radius  $r$  is the norm of  $\mathbf{x}$ . Fig. 5 plots the retinal filter  $\phi(r, t)$  as a function of  $r = \|\mathbf{x}\|_2$  for two different cases related to the temporal filters  $R_c(t)$  and  $R_s(t)$  which are studied in the following subsection. For simplicity, it is assumed that  $r \in \mathbb{R}$  and  $\phi(r, t)$  is symmetric around  $r = 0$  for all  $t$ . The parameters have been tuned according to neuroscientific results [3] which approximate the retinal spectrum and the speed of the retinal processing. Since the retina-inspired filter is a 2D spatially symmetric function, its spectrum is also symmetric. This spectrum, denoted  $\hat{\phi}(\omega, t)$  where  $\omega$  is the spatial angular frequency related to  $r$ , is shown in Fig. 5.

#### B. Spatiotemporal Behavior and Convergence

We are interested in studying the two temporal filters  $R_c(t)$  and  $R_s(t)$  which are responsible for the spatiotemporal behavior of the retina-inspired filter. First of all, we calculate their closed-form model in Proposition 2. This proposition is based on the following lemma.

**Lemma 1.** Assume  $t \geq 0$ , then

$$\begin{aligned} J_c(t) &= \int_0^t V(u) du \\ &= P_n(t) \exp\left(\frac{-t}{\tau_G}\right) + \alpha_c \exp\left(\frac{-t}{\tau_C}\right) + \gamma_c, \end{aligned} \quad (15)$$



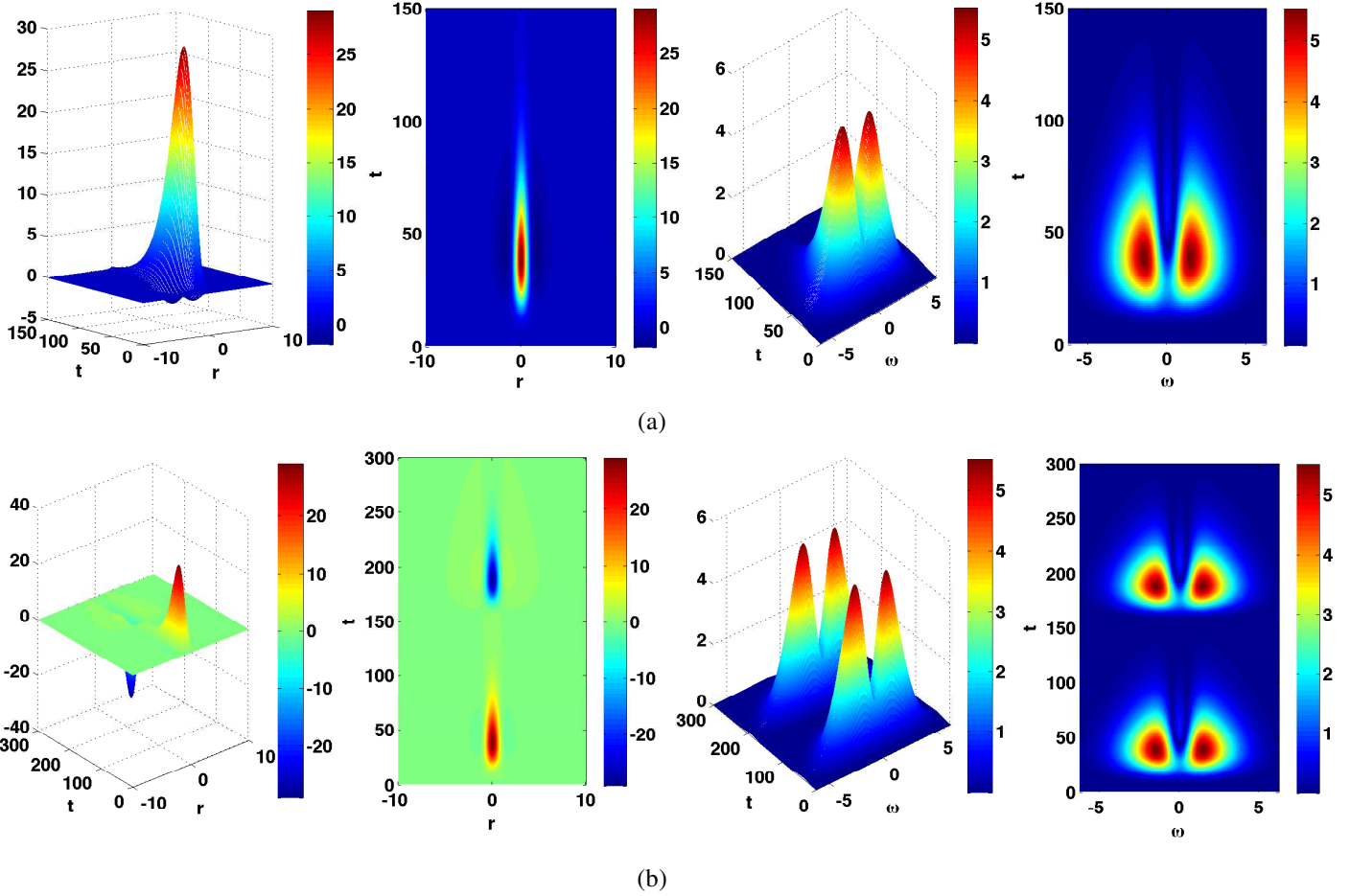


Fig. 5: (Left to right) The retina-inspired filter  $\phi(r, t)$  as a function of  $r \in \mathbb{R}$  and  $t \in \mathbb{R}$ . The top view of  $\phi(r, t)$ . The retina-inspired filter spectrum  $\hat{\phi}(\omega, t)$ . The top-view of  $|\hat{\phi}(\omega, t)|$ . (a) The top line corresponds to the case  $T = +\infty$ . (b) The bottom line stands for the case  $T < +\infty$  (Parameters:  $T = 150$  ms,  $\tau_C = 20$  ms,  $\tau_S = 4$  ms,  $\tau_G = 5$  ms,  $n = 5$ ,  $w_S = 1$ ,  $w_C = 1$ ,  $\sigma_c = 0.5$  and  $\sigma_s = 3\sigma_c$ ).

where  $P_n(t)$  is a polynomial function in  $t$  of order  $n$  and  $\alpha_c$  and  $\gamma_c$  are two reals, and

$$\begin{aligned} J_s(t) &= \int_0^t (V * E_{\tau_S})(u) du \\ &= Q_n(t) \exp\left(\frac{-t}{\tau_G}\right) + \alpha_s \exp\left(\frac{-t}{\tau_S}\right) \\ &\quad + \beta_s \exp\left(\frac{-t}{\tau_C}\right) + \gamma_s, \end{aligned} \quad (16)$$

where  $Q_n(t)$  is a polynomial function in  $t$  of order  $n$  and  $\alpha_s$ ,  $\beta_s$  and  $\gamma_s$  are some reals.

*Proof.* See the Appendix.  $\square$

**Proposition 2.** The temporal weights  $R_c(t)$  in (12) and  $R_s(t)$  in (13) satisfy:

$$R_c(t) = \begin{cases} J_c(t) & \text{if } 0 \leq t \leq T, \\ J_c(t) - J_c(t - T) & \text{if } T < t, \end{cases} \quad (17)$$

$$R_s(t) = \begin{cases} J_s(t) & \text{if } 0 \leq t \leq T, \\ J_s(t) - J_s(t - T) & \text{if } T < t, \end{cases} \quad (18)$$

where  $J_c(t)$  is given in (15) and  $J_s(t)$  is given in (16).

*Proof.* The proof is based on the fact that  $R_c(t) = J_c(t)$  for  $0 \leq t \leq T$  and

$$R_c(t) = \int_{t-T}^t V(u) du = \int_0^t V(u) du - \int_0^{t-T} V(u) du$$

for  $t > T$ . Lemma 1 is used to deduce (17). The same equalities hold for  $R_s(t)$ .  $\square$

The temporal weights  $R_c(t)$  and  $R_s(t)$  are illustrated in Fig. 6. The parameters in (a) and (b) have been tuned according to the parameters of Fig. 5 for  $T = +\infty$  and  $T < +\infty$  respectively. We should note that their shapes are very similar except that the surround temporal filter  $R_s(t)$  appears with a small delay  $E_{\tau_S}(t)$  with respect to the center one. There is a high impact of the above characteristic on the spatiotemporal evolution of the filter. The delay  $E_{\tau_S}(t)$  is crucially important because for the very first few milliseconds, while  $R_s(t)$  does not yet exist, the second term of the WDoG is zero. As a result, at the very beginning, the retina-inspired filter is a pure Gaussian with a very low amplitude since it is weighted by

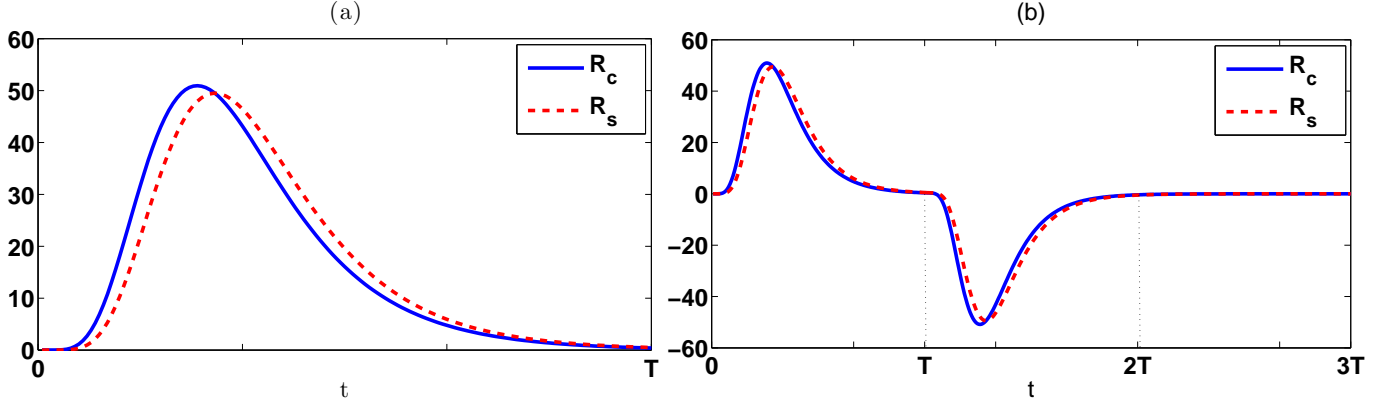


Fig. 6: Temporal filters  $R_c(t)$  and  $R_s(t)$ . The convergence of the filters depends on the value of time  $T$ . Subplot (a) corresponds to the case  $T = +\infty$ , while subplot (b) corresponds to the case  $T < +\infty$ .

$R_c(t)$ . Finally, we can note that  $R_c(t)$  and  $R_s(t)$  converge to a constant value. Hence,  $\phi(\mathbf{x}, t)$  also converges as  $t \rightarrow +\infty$ . This convergence is established in the following proposition.

**Proposition 3.** *The filter  $\phi(\mathbf{x}, t)$  is a continuous and infinitely differential function over  $\mathbb{R}^2 \times \mathbb{R}$  such that  $\phi(\mathbf{x}, 0) = 0$  for all  $\mathbf{x} \in \mathbb{R}^2$ . If  $T = +\infty$ , then  $\phi(\mathbf{x}, t)$  converges uniformly toward  $\phi(\mathbf{x})$  where  $\phi(\mathbf{x})$  is the WDoG filter:*

$$\phi(\mathbf{x}) = w_c \gamma_c G_{\sigma_c}(\mathbf{x}) - w_s \gamma_s G_{\sigma_s}(\mathbf{x}), \quad (19)$$

with  $\gamma_c$  and  $\gamma_s$  defined in (15)-(16). If  $T < +\infty$ , the filter vanishes uniformly as  $t \rightarrow +\infty$ :

$$\lim_{t \rightarrow +\infty} \sup_{\mathbf{x} \in \mathbb{R}^2} |\phi(\mathbf{x}, t)| = 0. \quad (20)$$

*Proof.* The uniform convergence (20) results from the definition of  $\phi(\mathbf{x}, t)$  in (11) since  $G_{\sigma_c}(\mathbf{x})$  and  $G_{\sigma_s}(\mathbf{x})$  are bounded and  $R_c(t)$  and  $R_s(t)$  converges to 0 according to Proposition 2 and Lemma 1. When  $T = +\infty$ ,  $R_c(t)$  and  $R_s(t)$  converge, respectively, to  $\gamma_c$  and  $\gamma_s$  as  $t \rightarrow +\infty$ . Hence,  $\phi(\mathbf{x}, t)$  converges uniformly toward  $\phi(\mathbf{x})$ .  $\square$

Figure 5 depicts the two different cases of Proposition 3. A brief discussion about this proposition is that, while time  $T$  increases to infinity, the retina-inspired filter turns into a static WDoG. This is totally concurrent with the neuroscientific assumptions about the time limits of the visual system. Neuroscientists have proposed that the objects categorization of a single image which is propagated from the retina to the brain needs approximately 100 ms [34] before the next image is processed. Recent studies prove that a simple comprehension lasts approximately 13 ms [35]. In any case, there exists a time  $t_c$  when, even if the photoreceptors will continue capturing the same signal, all the necessary information which need to be processed has already been transmitted to the brain. From a theoretical point of view, the existence of  $t_c$  is deduced from the uniform convergence established in Proposition 3. In fact, given  $\varepsilon > 0$ , the time  $t_c = t_c(\varepsilon)$  can be defined when the uniform convergence is achieved up to  $\varepsilon$ .

#### IV. WEIGHTED DOG ANALYSIS

This section aims to study the WDoG filter, i.e., to approximate its spatial response and its frequency response. Without any loss of generality, a WDoG is defined by:

$$\varphi(\mathbf{x}) = a G_{\sigma_a}(\mathbf{x}) - b G_{\sigma_b}(\mathbf{x}) \quad (21)$$

where  $a, b > 0$  and  $\sigma_b^2 > \sigma_a^2$ . The retina-inspired filter  $\phi(\mathbf{x}, t)$  consists of a group of WDoG with the coefficients

$$a = w_c R_c(t) = a(t), \quad b = w_s R_s(t) = b(t), \quad (22)$$

that are time dependent,  $\sigma_a = \sigma_c$  and  $\sigma_b = \sigma_s$ . Since the WDoG is symmetric, we define the WDoG according to the radial coordinate  $r = \|\mathbf{x}\|_2$ :

$$\varphi(r) = a G_{\sigma_a}(r) - b G_{\sigma_b}(r). \quad (23)$$

Since  $\varphi(r)$  is symmetric around 0, we assume that  $r \in \mathbb{R}$  (and not only to  $\mathbb{R}^+$ ) to ensure a better legibility of the results and to study more easily the spectrum of the WDoG. According to the couple  $(a, b)$ , the WDoG has eight shapes which are depicted in Fig. 7. The parameter  $b = 1$  is fixed. According to the value of  $a \in \{0.1, 0.3, 0.4, 0.7, 0.9, 1.1, 2, 4\}$ , the WDoG is either a lowpass filter (cases L1, L2 or L3) or a bandpass filter (case BP) or a mixed lowpass/bandpass filter (case LB). The conventional DoG filter ( $a = b = 1$ ) is a special case of BP. When  $a$  is changing, the bandwidth is also changing but the WDoG is always constrained to one of these eight shapes, corresponding to five behaviors. This section studies these behaviors in the space domain and in the frequency domain.

##### A. WDoG in Space Domain

Let us study the variations of  $\varphi(r)$ . The first derivative of the WDoG, which is differentiable for all  $r \in \mathbb{R}$ , is given by

$$\varphi'(r) = r \left( -\frac{a}{\sigma_a^2} G_{\sigma_a}(r) + \frac{b}{\sigma_b^2} G_{\sigma_b}(r) \right). \quad (24)$$

Let

$$\gamma = \gamma(a, b) = \frac{b \sigma_a^4}{a \sigma_b^4}. \quad (25)$$

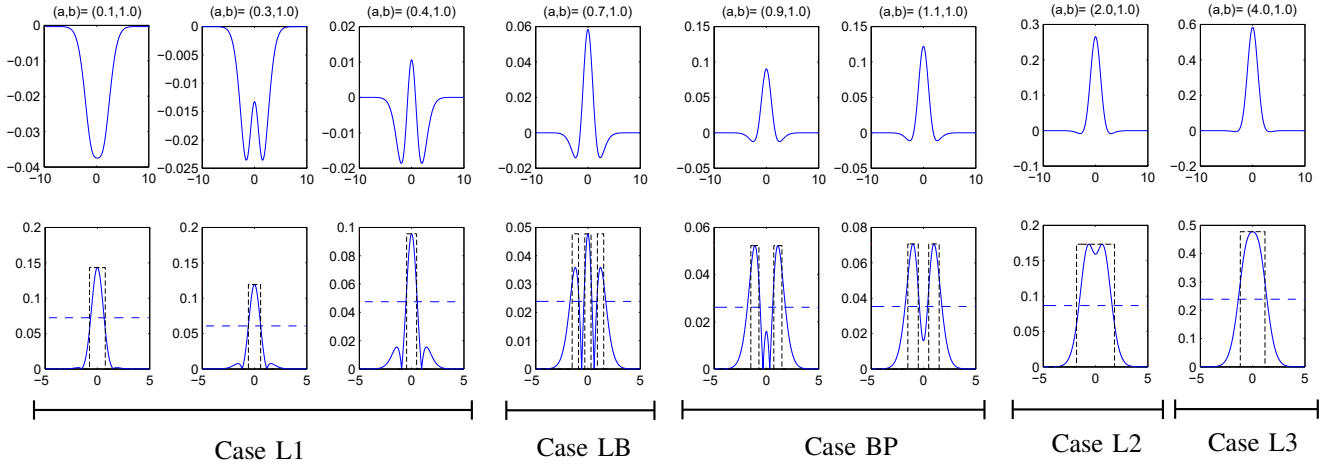


Fig. 7: Eight typical shapes of the WDoG  $\varphi(r)$  (first row) and the WDoG spectrum modulus  $|\hat{\varphi}(\omega)|$  (second row) w.r.t. some values of the couple  $(a, b)$ . The dotted line represents the half of the maximum value of the spectrum modulus. The rectangles in dashed line represent the approximated bandwidth of the spectrum. There are five specific cases: 1) lowpass L1, 2) lowpass/bandpass LB, 3) bandpass BP, 4) lowpass L2 and 5) lowpass L3.

A short analysis of the roots of  $\varphi'(r)$  shows that two cases occur: i) if  $\gamma \geq 1$ , there is only one root  $r = 0$ , ii) if  $\gamma < 1$ , there are three roots  $r = 0, r_1 > 0$  and  $-r_1$  with

$$r_1 = \sigma_a \sigma_b \sqrt{\frac{2 \ln(\gamma)}{\sigma_b^2 - \sigma_a^2}}. \quad (26)$$

As an illustration, case i) corresponds to the first left curve in Fig. 7 and case ii) corresponds to all the other curves. It is then easy to determine the positive intervals and negative intervals of  $\varphi'(r)$  and, hence, to determine when  $\varphi(r)$  is increasing or decreasing.

**Proposition 4.** *If  $\gamma \geq 1$ ,  $\varphi(r)$  is negative for all  $r \in \mathbb{R}$  and its minimum is*

$$\varphi_0 = \min_{r \in \mathbb{R}} \varphi(r) = \varphi(0) = \frac{a\sigma_b^2 - b\sigma_a^2}{2\pi\sigma_c^2\sigma_s^2} < 0. \quad (27)$$

*Otherwise, if  $\gamma < 1$ , then  $\varphi(r)$  is negative and decreasing over  $[-\infty, -r_1]$ , increasing over  $[-r_1, 0]$ , decreasing over  $[0, r_1]$  and finally increasing and negative over  $[r_1, +\infty]$ . The global minima are  $-r_1$  and  $r_1$ .*

*Proof.* The proof is straightforward by studying the sign of  $\varphi'(r)$  over the intervals given in the proposition.  $\square$

The Full Width Half Maximum (FWHM) response of the symmetric WDoG is given by the interval  $[-r_C, r_C]$  where  $r_C$  satisfies:

$$\varphi(r_C) = \frac{\max_{r \in \mathbb{R}} \varphi(r)}{2}. \quad (28)$$

The FWHM is not relevant for case L1 (see Fig. 7) because the shape of the WDoG is significantly different from a peak. The FWHM is relevant for the other cases when the peak at  $r = 0$  is sufficiently large, i.e., when  $\varphi(0) > -2\varphi(r_1)$ . The exact calculation of FWHM is difficult. Hence, we approximate  $r_C$  in (28) using the following method which is proposed in [36]. We first calculate the maximum response. Then, we compute a straight line from each maximum response to the first intercept

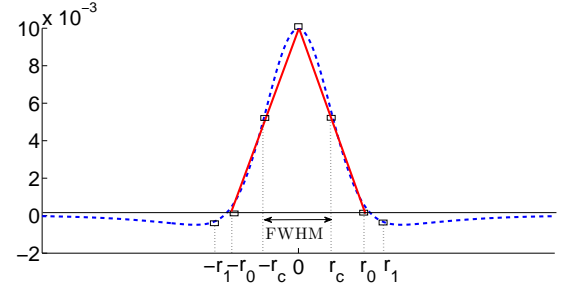


Fig. 8: Approximate FWHM interval of a WDoG based on a triangle approximation.

with the  $r$ -axis (see Fig. 8 red solid line). We compute half of this line and we approximate its correspondent value. The resulting FWHM is illustrated in Fig. 8. A straightforward calculation shows that the first intercept of  $\varphi(r)$  with the  $r$ -axis is

$$r_0 = \sigma_a \sigma_b \sqrt{\frac{2 \ln(\frac{a\sigma_b^2}{b\sigma_a^2})}{\sigma_b^2 - \sigma_a^2}}. \quad (29)$$

where  $b\sigma_a^2 > a\sigma_b^2$  since  $\varphi(0) > 0$ . It follows that  $r_C \approx \frac{r_0}{2}$  by using the triangle approximation.

### B. WDoG in Frequency Domain

The Fourier transform of the WDoG in (23) is

$$\hat{\varphi}(\omega) = a\hat{G}_{\sigma_a}(\omega) - b\hat{G}_{\sigma_b}(\omega) \quad (30)$$

where

$$\hat{G}_{\sigma}(\omega) = \frac{1}{2\pi} \exp\left(-\frac{\omega^2 \sigma^2}{2}\right), \quad (31)$$

and  $\omega \in \mathbb{R}$  denotes the spatial angular frequency associated to  $r$ . The WDoG bandwidth  $B = B(a, b)$  refers to the frequency range in which the spectrum  $\hat{\varphi}(\omega)$  is above a threshold value. The threshold value is defined relative to the maximum value



and the points where the spectrum is half its maximum value, i.e., we need to find all the solutions  $\bar{\omega}$  of

$$\hat{\varphi}(\bar{\omega}) = \frac{\max_{\omega \in \mathbb{R}} \hat{\varphi}(\omega)}{2}. \quad (32)$$

For each case identified in Fig. 7, the derivation of a closed form expression of the bandwidth is very tricky. Hence, we propose some simple approximations. For cases L1, LB, BP, each part of the bandwidth is approximated by a triangle. A triangle is formed by taking the zero position, the maximum position and double the turning point as already illustrated in Fig. 8. For cases L2 and L3, a better approximation is obtained by considering that the WDoG is almost equivalent to a Gaussian function.

Let us determine the zero position and the maximum position of the WDoG. The solutions of  $\hat{\varphi}(\omega) = 0$  are the two opposite roots  $\omega_0$  and  $-\omega_0$  with

$$\omega_0 = \sqrt{\frac{2}{\sigma_b^2 - \sigma_a^2} \ln\left(\frac{b}{a}\right)} \quad (33)$$

when  $b > a$ . Otherwise, when  $b \leq a$ , the WDoG is always positive with the single root  $\omega = 0$  in the special case  $a = b$ . Since the WDoG is differentiable, the extrema are the solutions of  $\hat{\varphi}'(\omega) = 0$  where  $\hat{\varphi}'(\omega)$  is the first derivative of  $\hat{\varphi}(\omega)$ . A short calculation shows that

$$\hat{\varphi}'(\omega) = \frac{\omega}{2\pi} \left( -a\sigma_a^2 \exp\left(-\frac{\omega^2 \sigma_a^2}{2}\right) + b\sigma_b^2 \exp\left(-\frac{\omega^2 \sigma_b^2}{2}\right) \right).$$

A first extrema is  $\omega = 0$ . The other extrema are the solutions of

$$\exp\left(\frac{\omega^2(\sigma_b^2 - \sigma_a^2)}{2}\right) = \frac{b\sigma_b^2}{a\sigma_a^2} = \varrho(a, b). \quad (34)$$

If  $\varrho(a, b) \leq 1$ , there is no other extrema. Otherwise, there are two opposite extrema  $\omega_1 > 0$  and  $-\omega_1$  where

$$\omega_1 = \sqrt{\frac{2}{\sigma_b^2 - \sigma_a^2} \ln(\varrho(a, b))}. \quad (35)$$

From (34), it is clear that  $\hat{\varphi}'(\omega) > 0$  for  $0 < \omega < \omega_1$  and  $\hat{\varphi}'(\omega) > 0$  for  $\omega_1 < \omega$ . Since  $\lim_{\omega \rightarrow \pm\infty} \hat{\varphi}(\omega) = 0$ , the frequency  $\omega_1$ , if it exists, is a global maximum. We can also show that  $-\omega_1$  is a global maximum. We obtain the following proposition.

**Proposition 5.** *If  $\varrho(a, b) > 1$ , then the maximum is given by*

$$\begin{aligned} \max_{\omega \in \mathbb{R}} \hat{\varphi}(\omega) &= \hat{\varphi}_1 = \hat{\varphi}(\omega_1) = \hat{\varphi}(-\omega_1) \\ &= \frac{a(\sigma_b^2 - \sigma_a^2)}{2\pi} \left( \frac{1}{\varrho(a, b)} \right)^{\frac{\sigma_a^2}{\sigma_b^2 - \sigma_a^2}} \end{aligned} \quad (36)$$

with  $\omega_1$  given in (35). Otherwise, if  $\varrho(a, b) \leq 1$ , then

$$\max_{\omega \in \mathbb{R}} \hat{\varphi}(\omega) = \hat{\varphi}_0 = \hat{\varphi}(0) = \frac{a - b}{2\pi} > 0.$$

*Proof.* If  $\varrho(a, b) \leq 1$ , there is only one extrema  $\omega = 0$ . Since  $\hat{\varphi}'(\omega)$  has the sign of  $-\omega$ ,  $\hat{\varphi}$  is increasing when  $\omega < 0$  and decreasing when  $\omega > 0$ . Hence, 0 is a positive global maximum. If  $\varrho(a, b) > 1$ , it has been already shown that the maximum are located at  $\omega_1$  and  $-\omega_1$ .

The maximum value in (36) is obtained by inserting  $\omega_1$ , given in (35), in (30).  $\square$

According to the couple  $(a, b)$ , the WDoG has three possible behaviors: lowpass, bandpass or lowpass/bandpass. Hence, the total bandwidth  $B = B(a, b)$ , including negative and positive frequencies, can have one of the three following forms:

- Lowpass:  $B = [-\omega_H, \omega_H]$  with  $\omega_H > 0$ ,
- Bandpass:  $B = [-\omega_H, -\omega_L] \cup [\omega_L, \omega_H]$  with  $0 < \omega_L < \omega_H$ ,
- Lowpass/bandpass:  $B = [-\omega_C, \omega_C] \cup [-\omega_H, -\omega_L] \cup [\omega_L, \omega_H]$  with  $0 < \omega_C < \omega_L < \omega_H$ .

The following proposition gives the bandwidth  $B(a, b)$  with respect to the couple  $(a, b)$ .

**Proposition 6.** *According to the value of  $(a, b)$ , the WDoG  $\hat{\varphi}(\omega)$  satisfies one of the following cases:*

- 1) *If  $\varrho > 1$  and  $|\hat{\varphi}_0| \geq \hat{\varphi}_1$* 
  - a) **Case L1:** *if  $|\hat{\varphi}_0| \geq 2\hat{\varphi}_1$ , then  $\hat{\varphi}(\omega)$  is lowpass with*

$$\omega_H \simeq \frac{\omega_0}{2},$$
  - b) **Case LB:** *if  $|\hat{\varphi}_0| < 2\hat{\varphi}_1$ , then  $\hat{\varphi}(\omega)$  is lowpass/bandpass with*

$$\omega_C \simeq \frac{\omega_0}{2}, \omega_L \simeq \frac{\omega_0 + \omega_1}{2}, \omega_H \simeq \frac{\omega_0 + 3\omega_1}{2},$$
- 2) *If  $\varrho > 1$  and  $|\hat{\varphi}_0| < \hat{\varphi}_1$* 
  - a) **Case BP:** *if  $|\hat{\varphi}_0| \leq \frac{\hat{\varphi}_1}{2}$ , then  $\hat{\varphi}(\omega)$  is bandpass with*

$$\omega_L \simeq \frac{\omega_0 + \omega_1}{2}, \omega_H \simeq \frac{\omega_0 + 3\omega_1}{2},$$
  - b) **Case L2:** *if  $|\hat{\varphi}_0| > \frac{\hat{\varphi}_1}{2}$ , then  $\hat{\varphi}(\omega)$  is lowpass with*

$$\omega_H \simeq \omega_1 + \frac{\sqrt{2 \ln(2)}}{\sigma_a},$$
- 3) *If  $\varrho \leq 1$ , which corresponds to Case L3, then  $\hat{\varphi}(\omega)$  is lowpass with*

$$\omega_H \simeq \frac{\sqrt{2 \ln(2)}}{\sigma_a}.$$

*Proof.* The behavior of the filter depends on the maximum values  $\hat{\varphi}_0$  or  $\hat{\varphi}_1$ , the maximum position and the zero position. In case of L1, LB and BP, each part of the bandwidth is approximated by a triangle. A short calculation gives the bounds of the triangle. In case of L2, we use two Gaussians to approximate the central part of the spectrum. The bandwidth of a Gaussian function with zero mean and standard deviation  $\sigma$  is given by

$$B_\sigma = \left[ -\sigma \sqrt{2 \ln(2)}, \sigma \sqrt{2 \ln(2)} \right].$$

In case of L3, we use a single Gaussian approximation.  $\square$

The retina-inspired filter is filtering the input image  $f(x)$  by using a WDoG varying in time. By controlling the trajectory  $(a(t), b(t))$ , this filter is able to explore the frequency spectrum of the input image.

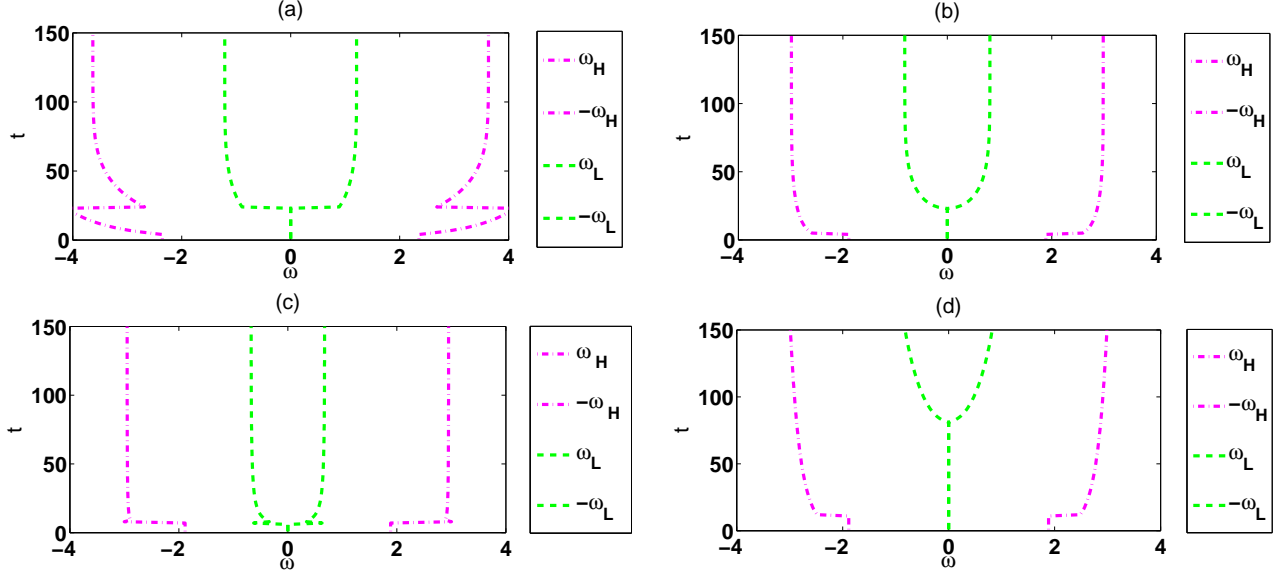


Fig. 9: For bio-plausible parameters described in Fig. 5 we compute in (a) the approximation of the evolution of the bandwidth with respect to time which is computed according to Proposition 6 and in (b) the exact evolution of the bandwidth with respect to time which is numerically computed. (c). The exact solution of the bandwidth which behaves as lowpass for longer time ( $\tau_C = 30$  ms,  $\tau_S = 20$  ms,  $\tau_G = 15$  ms,  $n = 5$ ). (b) The exact solution of the bandwidth which behaves as bandpass ( $\tau_C = 40$  ms,  $\tau_S = 0.5$  ms,  $\tau_G = 9$  ms).

## V. NUMERICAL RESULTS

This section aims to compute the retina-inspired decomposition tuned with respect to bio-plausible parameters given in [16] and described in Fig. 5. One could adjust these parameters and achieve non bio-plausible WDoG filters which are maybe more adaptable for different research topics. The tuning of these parameters is out of the scope of this work but we are going to test some non bio-plausible parameters and discuss the results. In addition, concerning the two different case when  $T = +\infty$  and  $T < +\infty$ , one could notice according to Fig. 6 that the second one is separated into three groups for  $t \in [0, T]$ ,  $t \in (T, 2T]$  and  $t > 2T$ . Concerning the first group, the retina-inspired filter would evolve exactly in the case was as in case  $T = +\infty$ . The second group is the inverse of the first group and the third group corresponds to the zero value because of the zero convergence. According to Fig. 5 it is obvious that at least for bio-plausible parameters, the two first groups have exactly the same spectrum. As a result, we concentrate our analysis on the first group which is identical to  $T = +\infty$ .

We are able now to compute the bandwidth  $B = B(a, b)$  of the filter as a function of time using Proposition 6 which is illustrated in Fig. 9 (a). The approximation consists of 3 different case which appear progressively in time: for  $1 \leq t \leq 5$  ms the bandwidth is approximated according to L3, for  $5 < t \leq 23$  ms it follows L2 and finally, for  $t > 23$  ms it belongs to BP. The exact bandwidth is computed by solving the equation (32) using the Matlab<sup>TM</sup> *fsolve* function and it is depicted in Fig. 9 (b). Concerning the exact solution one should notice that the bandwidth evolves in time in a very similar way depicted that the filter behaves as a lowpass for  $1 \leq t \leq 23$  ms while it becomes bandpass for  $t > 23$  ms. Comparing these two plots, we confirm the accuracy of Proposition 6 is not perfect

because of the triangle approximation. This approximation of the DoG filter has already been used in previous works getting similar accuracy results [36]. It would be also important to be mentioned that the approximation is not continuous but it has been proposed for each case separately. This is the reason why some discontinuities appear. The second row of Fig. 9 comes to remind the reader that if one chooses different set of parameters, not only the behavior of the filter but also the evolution of the bandwidth with respect to time will change.

In Fig. 10, we represent the filtering results while the retina-inspired filter is applied to an image for different parameters of  $w_c$  and  $w_s$ . We used a gray-scale image of size  $n = 512 \times 512$  pixels. We have decided to illustrate only 5 decomposition layers for each experiment. The evolution of the retina-inspired filter is according to the bandwidth of Fig. 9 (b). The first column of Fig. 10 corresponds to the bio-plausible parameters where  $w_c = w_s = 1$ . The middle and the left columns correspond to non bio-plausible parameters but very close to the ones of the left column. It is interesting that while  $w_c$  decreases there is a stronger categorization of the L2, L3, LB and BP cases. Another important remark is related to the scale of the top pictures which correspond to the first decomposition layer. Apparently, this layer seems to be identical to the original. However, its scale is very low (i.e.  $10^{-13}$ ) which means that in the presence of noise or after quantization all this information will be completely lost.

## VI. CONCLUSION

This paper has introduced a non-separable spatiotemporal retina-inspired filter based on a realistic model of the OPL. This filter is a groundbreaking analysis of neuroscience for image processing. It has been established that this filter is

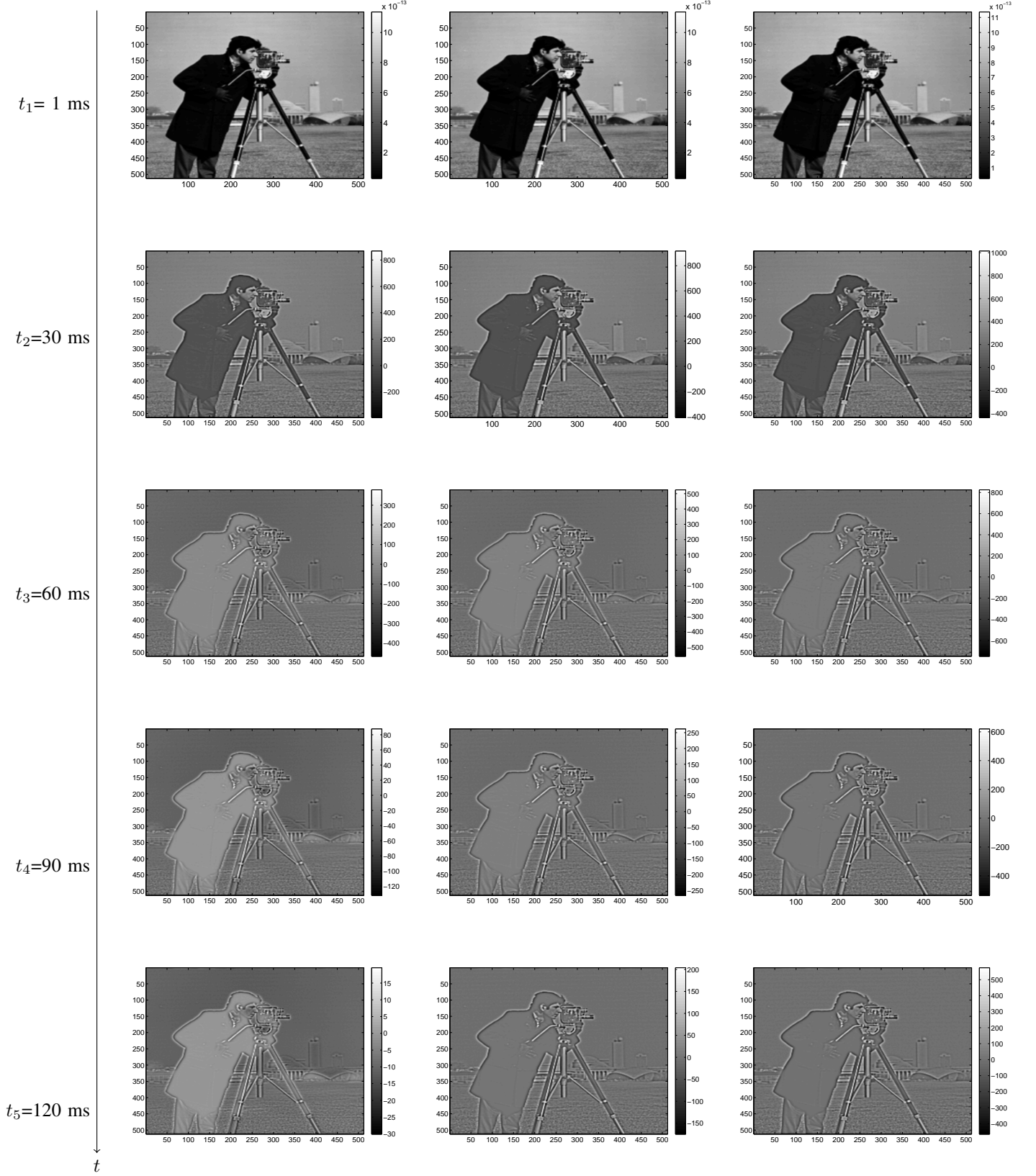


Fig. 10: Decomposition of an image using the retina-inspired non-separable spatiotemporal filter. Left:  $w_c = w_s = 1$ . Middle:  $w_c = 0.9$  and  $w_s = 1$ . Right:  $w_c = 0.7$  and  $w_s = 1$ .

a group of time-varying WDoGs when it is applied to a still image flashed for a given time. Hence, it enables the extraction of different kind of data while time increases. As underlined in the introduction, the retina-inspired filter is a great improvement of other bio-inspired filters which are simpler and not much as accurate neuroscientific approximation of the retina. Hence, this is certainly of a great interest for the image processing field.

There are numerous applications for our retinal-inspired filter, the most important of which is image compression. We believe that the retina-inspired transform is the first step to build novel retina-inspired codecs. It seems that even if image and video codecs are still two of the most attractive research topics, they have reached their limits: even if the bi-rate efficiently decreases, the computational cost dramatically increases [37]. The HVS has low power consumption while at the same time its visual perception is very high. Some models have been already proposed to mimic the retinal coding like the Rank Order Coding (ROC) [3], [38], [39]. However, the filtering bank used by ROC was very simple. As a result, it would be interesting to extend these works by using our retina-inspired model. In addition, it would be also relevant to study a full bio-plausible neural coding architecture like [15], [40] based on our model.

## APPENDIX

This appendix proves Lemma 1. It is shown that both  $J_c(t)$  and  $J_s(t)$  can be defined in a closed form as polynomial functions which are attenuated by exponential ones. These functions will be essential to calculate  $R_c(t)$  and  $R_s(t)$ . The calculation of  $J_c(t)$  and  $J_s(t)$  are based on the following lemma whose proof is straightforward.

**Lemma 2.** *Let  $\omega$  a real value,  $t \geq 0$  and  $n$  a positive integer. Using an integration by parts, we obtain the following equality:*

$$\int_0^t u^n \exp(-\omega u) du = P_n(t) \exp(-\omega t) + c,$$

where

$$P_n(t) = \sum_{k=0}^n -\frac{n!}{(n-k)! \omega^{k+1}} t^{n-k}$$

is a polynomial function in  $t$  of order  $n$  whose coefficients depend on  $\omega$  and  $c$  is a constant value.

### A. Closed-form of $J_c(t)$

Assume that  $0 \leq t \leq T$ . It follows that:

$$J_c(t) = \int_{u=0}^t V(u) du,$$

which yields

$$\begin{aligned} J_c(t) &= \int_{u=0}^t E_{\tau_G, n} {}^t (\delta_0 - w_C E_{\tau_C})(u) du \\ &= \int_{u=0}^t E_{\tau_G, n}(u) du - w_C \int_{u=0}^t E_{\tau_G, n} {}^t E_{\tau_C}(u) du. \end{aligned}$$

The definition of the gamma and exponential filters yields:

$$\begin{aligned} J_c(t) &= \frac{1}{\tau_G^{n+1}} \int_{u=0}^t u^n \exp(-au) du \\ &- \frac{w_C}{\tau_G^{n+1} \tau_C} \left( \int_{u=0}^t \exp\left(\frac{-u}{\tau_C}\right) \int_{v=0}^u v^n \exp(-bv) dv du \right). \end{aligned}$$

where  $a = \frac{1}{\tau_G}$ ,  $b = \frac{\tau_C - \tau_G}{\tau_G \tau_C}$ . Using Lemma 2, we get

$$\begin{aligned} J_c(t) &= \sum_{k=0}^n -\frac{(n!)}{(n-k)! a^{k+1} \tau_G^{n+1}} t^{n-k} \exp(-at) \\ &- \sum_{k=0}^n \sum_{l=0}^m \frac{n! w_C}{(m-l)! b^{k+1} a^{l+1} \tau_G^{n+1} \tau_C} t^{m-l} \exp(-at) \\ &+ \frac{n! w_C}{b^{n+1} \tau_G^{n+1}} \exp\left(\frac{-t}{\tau_C}\right) \\ &+ \sum_{k=0}^n \frac{n! w_C}{b^{k+1} a^{m+1} \tau_G^{n+1} \tau_C} - \frac{n! w_C}{b^{n+1} \tau_G^{n+1}} + n! \end{aligned}$$

where  $m = n - k$ . Finally,

$$J_c(t) = P_n(t) \exp\left(\frac{-t}{\tau_G}\right) + \alpha_c \exp\left(\frac{-t}{\tau_C}\right) + \gamma_c$$

where  $P_n(t)$  is a polynomial function in  $t$  of order  $n$  and  $\alpha_c$  and  $\gamma_c$  are two reals.

### B. Closed-form of $J_s(t)$

The method used to calculate  $J_c(t)$  can be applied to  $J_s(t)$ . This leads to:

$$\begin{aligned} J_s(t) &= \int_{u=0}^t (V {}^t E_{\tau_S})(u) du \\ &= \int_{u=0}^t E_{\tau_G, n} {}^t (\delta_0 - w_C E_{\tau_C}) {}^t E_{\tau_S}(u) du \\ &= \int_{u=0}^t E_{\tau_G, n} {}^t E_{\tau_S}(u) du \\ &- w_C \int_{u=0}^t E_{\tau_G, n} {}^t E_{\tau_C} {}^t E_{\tau_S}(u) du. \end{aligned}$$

Using Lemma 2, we get

$$\begin{aligned} J_s(t) &= \frac{1}{\tau_G^{n+1} \tau_S} \left( \sum_{k=0}^n \sum_{l=0}^m \frac{n! t^{m-l} \exp(-at)}{g^{k+1} a^{l+1} (m-l)!} \right. \\ &- \sum_{k=0}^n \frac{n!}{g^{k+1} a^{m+1}} + \frac{n! \tau_S}{g^{n+1}} \left( 1 - \exp\left(\frac{-t}{\tau_S}\right) \right) \Big) \\ &- \frac{w_C}{\tau_G^{n+1} \tau_C \tau_S} \left( \sum_{k=0}^n \sum_{l=0}^m \sum_{r=0}^p -\frac{n! t^{p-r} \exp(-at)}{(p-r)! b^{k+1} g^{l+1} a^{r+1}} \right. \\ &+ \sum_{k=0}^n \sum_{l=0}^m \frac{n!}{b^{k+1} g^{l+1} a^{p+1}} \\ &- \sum_{k=0}^n \frac{n! \tau_S}{b^{k+1} g^{m+1}} \left( 1 - \exp\left(\frac{-t}{\tau_S}\right) \right) \Big) \end{aligned}$$

$$+ \frac{n! \tau_S}{b^{n+1} \phi} \left( 1 - \exp \left( \frac{-t}{\tau_S} \right) \right) - \frac{n! \tau_C}{b^{n+1} \phi} \left( 1 - \exp \left( \frac{-t}{\tau_C} \right) \right),$$

with the variables

$$g = \frac{\tau_S - \tau_G}{\tau_G \tau_S}, \quad \phi = \frac{\tau_S - \tau_C}{\tau_C \tau_S},$$

$p = m - l$  and  $m = n - k$ . It follows that:

$$J_s(t) = Q_n(t) \exp \left( \frac{-t}{\tau_G} \right) + \alpha_s \exp \left( \frac{-t}{\tau_S} \right) + \beta_s \exp \left( \frac{-t}{\tau_C} \right) + \gamma_s$$

where  $Q_n(t)$  is a polynomial function in  $t$  of order  $n$  and  $\alpha_s$ ,  $\beta_s$  and  $\gamma_s$  are some reals. This ends the proof.

## REFERENCES

- [1] D. J. Field, "What is the goal of sensory coding?," *Neural Computation*, vol. 6, pp. 559–601, 1994.
- [2] G. DeAngelis D. Cai and R. Freeman, "Spatiotemporal receptive field organization in the Lateral Geniculate Nucleus of cats and kittens," *The American Physiological Society*, vol. 22, no. 3077, pp. 1045–1061, 1997.
- [3] K. Masmoudi, M. Antonini, and P. Kornprobst, "Frames for exact inversion of the rank order coder," *IEEE Transaction on Neural Networks*, vol. 23, no. 2, pp. 353–359, 2012.
- [4] H. Winnemöller, J. E. Kyprianidis, and S. C. Olsen, "XDoG: An eXtended difference-of-Gaussians compendium included advanced image stylization," *Computer and Graphics*, vol. 36, no. 6, pp. 720–753, 2012.
- [5] S. Pei and L. Chen, "Image quality assessment using human visual DOG model fused with random forest," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3282–2193, 2015.
- [6] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London. Series B, Biological Sciences*, vol. 207, no. 1167, pp. 187–217, 1980.
- [7] L. M. Kennedy and M. Basu, "Image enhancement using a human visual system model," *Pattern Recognition*, vol. 30, no. 12, pp. 2001–2014, 1997.
- [8] M. Basu, "Gaussian-based edge-detection methods—a survey," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 32, no. 2, pp. 252 – 260, 2002.
- [9] D. G. Lowe, "Object recognition from local scale-invariant features," *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, vol. 2, no. 1150 – 1157, 1999.
- [10] P. Burt and E. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Transactions on Communication*, vol. 31, no. 4, pp. 532–540, 1983.
- [11] E. Adelson, C. Adelson, J. Bergen, P. Burt, and J. Ogden, "Pyramid methods in image processing," *RCA engineer*, vol. 29, no. 6, pp. 33–41, 1984.
- [12] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions of Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, 1998.
- [13] S. Frintrop, "Traditional saliency reloaded: A good old model in new shape," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 82–89, 2015.
- [14] R. Young, "The gaussian derivative model for spatial-temporal vision: I. cortical model," *Spatial Vision*, vol. 14, no. 3,4, pp. 261–319, 2001.
- [15] K. Masmoudi, M. Antonini, and P. Kornprobst, "Streaming an image through the eye: The retina seen as a dithered scalable image coder," *Signal processing Image Communication*, vol. 28, no. 8, pp. 856–869, 2013.
- [16] A. Wohrer and P. Kornprobst, "Virtual retina: A biological retina model and simulator, with contrast gain control," *Journal of Computational Neuroscience*, vol. 26, no. 2, pp. 219–249, 2009.
- [17] E. Doutsis, L. Fillatre, M. Antonini, and J. Gaulmin, "Retina-inspired filtering for dynamic image coding," *IEEE International Conference in Image Processing (ICIP)*, pp. 3505 – 3509, 2015.
- [18] E. Doutsis, L. Fillatre, M. Antonini, and J. Gaulmin, "Event-based coding of images using a bio-inspired frame," *International Conference on Event-Based Control, Communication and Signal Processing (EBCCSP)*, 2015.
- [19] S. Kuffler, "Neurons in the retina: Organization, inhibition and excitation problems," *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 17, pp. 281–292, 1952.
- [20] D. Marr, *Vision*, The MIT Press, 1982.
- [21] D. Fleet, P. Hallett, and A. Jepson, "Spatiotemporal inseparability in early visual processing," *Biological Cybernetics*, vol. 52, pp. 153–164, 1985.
- [22] D. Hubel, "The visual cortex of the brain," *American Scientist the magazine of Sigma Xi, The Scientific Research Society*, vol. 209, no. 5, pp. 54–63, 1963.
- [23] M. Livingstone and D. Hubel, "Anatomy and physiology of a color system in the primate visual cortex," *Journal of Neuroscience*, vol. 4, no. 1, pp. 309–356, 1984.
- [24] R. Desimone and L. G. Ungerleider S. J. Schein, J. Moran, "Contour, color and shape analysis beyond the striate cortex," *Neural Basis of Visual Perception*, vol. 25, no. 3, pp. 441–452, 1985.
- [25] R. Masland, "The fundamental plan of the retina," *Natural Neuroscience*, vol. 4, no. 9, pp. 877–886, 2001.
- [26] H. Kolb, "How the retina works," *American Scientist the magazine of Sigma Xi, The Scientific Research Society*, vol. 91, pp. 28–35, 2004.
- [27] Bart M. ter Haar Romeny, *Front-End Vision and Multi-Scale Image Analysis*, computational imaging and vision The front-end visual system — the retina, pp. 153–165, Springer Netherlands, 2003.
- [28] D. VanEssen, C. Adelson, and D. Felleman, "Information processing in the primate visual system: An integrated systems perspective," *Science*, vol. 255, no. 5043, pp. 419–423, 2005.
- [29] A. Wohrer, P. Kornprobst, and M. Antonini, "Retinal filtering and image reconstruction," *Tech. Rep., Inria Research ReportRR- 6960*, 2009.
- [30] R. Masland, "Cell populations of the retina: The proctor lecture," *Investigative Ophthalmology and visual Science*, vol. 52, no. 7, pp. 4581–4591, 2011.
- [31] H. Ögmen and M. H. Herzog, "The geometry of visual perception: Retinotopic and nonretinotopic representations in the human visual system," *Proceedings of the IEEE*, vol. 98, no. 3, pp. 479–492, 2010.
- [32] Jeanny Hérault and Barthélemy Durette, *Modeling Visual Perception for Image Processing*, pp. 662–675, Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.
- [33] G. DeAngelis, I. Ohzawa, and R. Freeman, "Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. i. general characteristics and postnatal development," *Journal of Neurophysiology*, vol. 69, no. 4, pp. 1091– 1117, 1993.
- [34] H. Liu, Y. Agam, J. R. Madsen, and G. Kreiman, "Timing, timing, timing: Fast decoding of object information intracranial field potentials in human visual cortex," *Neuron*, vol. 62, pp. 281–290, 2009.
- [35] M. C. Potter, B. Wyble, C. E. Hagmann, and E. S. McCourt, "Detecting meaning in RSVP at 13ms per picture," *Attention, Perception & Psychophysics*, vol. 76, no. 2, pp. 270–279, 2014.
- [36] P. Birch, B. Mitra, N. M. Bangalore, S. Rehman, R. Young, and C. Chatwin, "Approximate bandpass and frequency response models of the difference of gaussian filter," *Optics Communications*, vol. 283, pp. 4942–4948, 2010.
- [37] D. Grois, D. Marpe, A. Mulyoff, and O. Hadar, "Performance comparison of h.265/mpeg-hevc, vp9, and h.264/mpeg-avc encoders," *30th Picture Coding Symposium 2013 (PCS 2013)*, December 2013.
- [38] S. J. Thorpe, "Spike arrival times: A highly efficient coding scheme for neural network," *Parallel Processing in neural Systems and Computers*, , no. 91-94, 1990.
- [39] S. J. Thorpe and J. Gautrais, "Rank Order Coding: A new coding scheme for rapid processing in neural network," *Computational Neuroscience: Trends in Research*, , no. 113-118, 1998.
- [40] A. A. Lazar and E. A. Pnevmatikakis, "Video time encoding machines," *IEEE Transaction on Neural Networks*, vol. 22, no. 3, pp. 461–473, 2011.